

#10 SV

www

PCT09

P.5

## RAW SEQUENCE LISTING

DATE: 11/02/2001

PATENT APPLICATION: US/09/856,298

TIME: 14:40:19

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11022001\I856298.raw

ENTERED

```

3 <110> APPLICANT: UEMURA, Hidetoshi
4   OKUI, Akira
5   KOMINAMI, Katsuya
6   YAMAGUCHI, Nozomi
7   MITSUI, Shinichi
9 <120> TITLE OF INVENTION: NOVEL SERINE PROTEASE BSSP4
11 <130> FILE REFERENCE: UEMURA=6
13 <140> CURRENT APPLICATION NUMBER: 09/856,298
14 <141> CURRENT FILING DATE: 2001-05-21
16 <150> PRIOR APPLICATION NUMBER: JP 10/347813
17 <151> PRIOR FILING DATE: 1998-11-20
19 <150> PRIOR APPLICATION NUMBER: PCT/JP99/06472
20 <151> PRIOR FILING DATE: 1999-11-19
22 <160> NUMBER OF SEQ ID NOS: 52
24 <170> SOFTWARE: PatentIn version 3.1
26 <210> SEQ ID NO: 1
27 <211> LENGTH: 1282
28 <212> TYPE: DNA
29 <213> ORGANISM: Homo sapiens
31 <220> FEATURE:
32 <221> NAME/KEY: CDS
33 <222> LOCATION: (4)..(954)
34 <223> OTHER INFORMATION:
37 <220> FEATURE:
38 <221> NAME/KEY: mat_peptide
39 <222> LOCATION: (151)..()
40 <223> OTHER INFORMATION:
43 <400> SEQUENCE: 1
44 gcc atg gtg gtt tct gga gcg ccc cca gcc ctg ggt ggg ggc tgt ctc      48
45   Met Val Val Ser Gly Ala Pro Pro Ala Leu Gly Gly Gly Cys Leu
46               -45                      -40                      -35
48 ggc acc ttc acc tcc ctg ctg ctg ctg gcg tcg aca gcc atc ctc aat      96
49 Gly Thr Phe Thr Ser Leu Leu Leu Leu Ala Ser Thr Ala Ile Leu Asn
50               -30                      -25                      -20
52 gcg gcc agg ata cct gtt ccc cca gcc tgt ggg aag ccc cag cag ctg      144
53 Ala Ala Arg Ile Pro Val Pro Pro Ala Cys Gly Lys Pro Gln Gln Leu
54               -15                      -10                      -5
56 aac cgg gtt gtg ggc ggc gag gac agc act gac agc gag tgg ccc tgg      192
57 Asn Arg Val Val Gly Gly Glu Asp Ser Thr Asp Ser Glu Trp Pro Trp
58   1 1          5          10
60 atc gtg agc atc cag aag aat ggg acc cac cac tgc gca ggt tct ctg      240
61 Ile Val Ser Ile Gln Lys Asn Gly Thr His His Cys Ala Gly Ser Leu
62 15          20          25          30
64 ctc acc agc cgc tgg gtg atc act gct gcc cac tgt ttc aag gac aac      288
65 Leu Thr Ser Arg Trp Val Ile Thr Ala Ala His Cys Phe Lys Asp Asn
66               35          40          45
68 ctg aac aaa cca tac ctg ttc tct gtg ctg ctg ggg gcc tgg cag ctg      336

```

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/856,298

DATE: 11/02/2001

TIME: 14:40:19

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11022001\I856298.raw

```

69 Leu Asn Lys Pro Tyr Leu Phe Ser Val Leu Leu Gly Ala Trp Gln Leu
70          50          55          60
72 ggg aac cct ggc tct cgg tcc cag aag gtg ggt gtt gcc tgg gtg gag      384
73 Gly Asn Pro Gly Ser Arg Ser Gln Lys Val Gly Val Ala Trp Val Glu
74          65          70          75
76 ccc cac cct gtg tat tcc tgg aag gaa ggt gcc tgt gca gac att gcc      432
77 Pro His Pro Val Tyr Ser Trp Lys Glu Gly Ala Cys Ala Asp Ile Ala
78          80          85          90
80 ctg gtg cgt ctc gag cgc tcc ata cag ttc tca gag cgg gtc ctg ccc      480
81 Leu Val Arg Leu Glu Arg Ser Ile Gln Phe Ser Glu Arg Val Leu Pro
82 95          100          105          110
84 atc tgc cta cct gat gcc tct atc cac ctc cct cca aac acc cac tgc      528
85 Ile Cys Leu Pro Asp Ala Ser Ile His Leu Pro Pro Asn Thr His Cys
86          115          120          125
88 tgg atc tca ggc tgg ggg agc atc caa gat gga gtt ccc ttg ccc cac      576
89 Trp Ile Ser Gly Trp Gly Ser Ile Gln Asp Gly Val Pro Leu Pro His
90          130          135          140
92 cct cag acc ctg cag aag ctg aag gtt cct atc atc gac tcg gaa gtc      624
93 Pro Gln Thr Leu Gln Lys Leu Lys Val Pro Ile Ile Asp Ser Glu Val
94          145          150          155
96 tgc agc cat ctg tac tgg cgg gga gca gga cag gga ccc atc act gag      672
97 Cys Ser His Leu Tyr Trp Arg Gly Ala Gly Gln Gly Pro Ile Thr Glu
98          160          165          170
100 gac atg ctg tgt gcc ggc tac ttg gag ggg gag cgg gat gct tgt ctg      720
101 Asp Met Leu Cys Ala Gly Tyr Leu Glu Gly Glu Arg Asp Ala Cys Leu
102 175          180          185          190
104 ggc gac tcc ggg ggc ccc ctc atg tgc cag gtg gac ggc gcc tgg ctg      768
105 Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Val Asp Gly Ala Trp Leu
106          195          200          205
108 ctg gcc ggc atc atc agc tgg ggc gag ggc tgt gcc gag cgc aac agg      816
109 Leu Ala Gly Ile Ile Ser Trp Gly Glu Gly Cys Ala Glu Arg Asn Arg
110          210          215          220
112 ccc ggg gtc tac atc agc ctc tct gcg cac cgc tcc tgg gtg gag aag      864
113 Pro Gly Val Tyr Ile Ser Leu Ser Ala His Arg Ser Trp Val Glu Lys
114          225          230          235
116 atc gtg caa ggg gtg cag ctc cgc ggg cgc gct cag ggg ggt ggg gcc      912
117 Ile Val Gln Gly Val Gln Leu Arg Gly Arg Ala Gln Gly Gly Gly Ala
118          240          245          250
120 ctc agg gca ccg agc cag ggc tct ggg gcc gcc gcg cgc tcc      954
121 Leu Arg Ala Pro Ser Gln Gly Ser Gly Ala Ala Arg Ser
122 255          260          265
124 tagggcgag cgggacgcgg ggcgcggatc tgaaaggcgg ccagatccac atctggatct      1014
126 ggatctgcgg cgccctcggg cggtttcccc cgccgtaaat aggcctcatct acctctacct      1074
128 ctggggggccc ggacggctgc tgcggaaagg aaaccccctc ccgacccgc ccgacggcct      1134
130 caggccccgc cctccaaggc atcaggcccc gcccaacggc ctcatgtccc cgccccacg      1194
132 acttcggcc ccgccccggg gccccagcgc ttttgtgtat ataaatgtta atgattttta      1254
134 taggtatttg taaccctgcc cacatatc                                1282
137 <210> SEQ ID NO: 2
138 <211> LENGTH: 317

```

## RAW SEQUENCE LISTING

DATE: 11/02/2001

PATENT APPLICATION: US/09/856,298

TIME: 14:40:19

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11022001\I856298.raw

```

139 <212> TYPE: PRT
140 <213> ORGANISM: Homo sapiens
142 <400> SEQUENCE: 2
144 Met Val Val Ser Gly Ala Pro Pro Ala Leu Gly Gly Gly Cys Leu Gly
145           -45                      -40                      -35
148 Thr Phe Thr Ser Leu Leu Leu Leu Ala Ser Thr Ala Ile Leu Asn Ala
149           -30                      -25                      -20
152 Ala Arg Ile Pro Val Pro Pro Ala Cys Gly Lys Pro Gln Gln Leu Asn
153           -15                      -10                      -5
156 Arg Val Val Gly Gly Glu Asp Ser Thr Asp Ser Glu Trp Pro Trp Ile
157 -1  1           5           10           15
160 Val Ser Ile Gln Lys Asn Gly Thr His His Cys Ala Gly Ser Leu Leu
161           20           25           30
164 Thr Ser Arg Trp Val Ile Thr Ala Ala His Cys Phe Lys Asp Asn Leu
165           35           40           45
168 Asn Lys Pro Tyr Leu Phe Ser Val Leu Leu Gly Ala Trp Gln Leu Gly
169           50           55           60
172 Asn Pro Gly Ser Arg Ser Gln Lys Val Gly Val Ala Trp Val Glu Pro
173           65           70           75
176 His Pro Val Tyr Ser Trp Lys Glu Gly Ala Cys Ala Asp Ile Ala Leu
177 80           85           90           95
180 Val Arg Leu Glu Arg Ser Ile Gln Phe Ser Glu Arg Val Leu Pro Ile
181           100          105          110
184 Cys Leu Pro Asp Ala Ser Ile His Leu Pro Pro Asn Thr His Cys Trp
185           115          120          125
188 Ile Ser Gly Trp Gly Ser Ile Gln Asp Gly Val Pro Leu Pro His Pro
189           130          135          140
192 Gln Thr Leu Gln Lys Leu Lys Val Pro Ile Ile Asp Ser Glu Val Cys
193           145          150          155
196 Ser His Leu Tyr Trp Arg Gly Ala Gly Gln Gly Pro Ile Thr Glu Asp
197 160           165          170          175
200 Met Leu Cys Ala Gly Tyr Leu Glu Gly Glu Arg Asp Ala Cys Leu Gly
201           180          185          190
204 Asp Ser Gly Gly Pro Leu Met Cys Gln Val Asp Gly Ala Trp Leu Leu
205           195          200          205
208 Ala Gly Ile Ile Ser Trp Gly Glu Gly Cys Ala Glu Arg Asn Arg Pro
209           210          215          220
212 Gly Val Tyr Ile Ser Leu Ser Ala His Arg Ser Trp Val Glu Lys Ile
213           225          230          235
216 Val Gln Gly Val Gln Leu Arg Gly Arg Ala Gln Gly Gly Gly Ala Leu
217 240           245          250          255
220 Arg Ala Pro Ser Gln Gly Ser Gly Ala Ala Ala Arg Ser
221           260          265
224 <210> SEQ ID NO: 3
225 <211> LENGTH: 1007
226 <212> TYPE: DNA
227 <213> ORGANISM: Homo sapiens
229 <220> FEATURE:
230 <221> NAME/KEY: CDS

```

## RAW SEQUENCE LISTING

DATE: 11/02/2001

PATENT APPLICATION: US/09/856,298

TIME: 14:40:19

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11022001\I856298.raw

```

231 <222> LOCATION: (4)..(960)
232 <223> OTHER INFORMATION:
235 <220> FEATURE:
236 <221> NAME/KEY: mat_peptide
237 <222> LOCATION: (151)..()
238 <223> OTHER INFORMATION:
241 <400> SEQUENCE: 3
242 gcc atg gtg gtt tct gga gcg ccc cca gcc ctg ggt ggg ggc tgt ctc      48
243     Met Val Val Ser Gly Ala Pro Pro Ala Leu Gly Gly Gly Cys Leu
244             -45             -40             -35
246 ggc acc ttc acc tcc ctg ctg ctg ctg gcg tcg aca gcc atc ctc aat      96
247 Gly Thr Phe Thr Ser Leu Leu Leu Leu Ala Ser Thr Ala Ile Leu Asn
248             -30             -25             -20
250 gcg gcc agg ata cct gtt ccc cca gcc tgt ggg aag ccc cag cag ctg      144
251 Ala Ala Arg Ile Pro Val Pro Pro Ala Cys Gly Lys Pro Gln Gln Leu
252             -15             -10             -5
254 aac cgg gtt gtg ggc ggc gag gac agc act gac agc gag tgg ccc tgg      192
255 Asn Arg Val Val Gly Gly Glu Asp Ser Thr Asp Ser Glu Trp Pro Trp
256     -1  1             5             10
258 atc gtg agc atc cag aag aat ggg acc cac cac tgc gca ggt tct ctg      240
259 Ile Val Ser Ile Gln Lys Asn Gly Thr His His Cys Ala Gly Ser Leu
260 15             20             25             30
262 ctc acc agc cgc tgg gtg atc act gct gcc cac tgt ttc aag gac aac      288
263 Leu Thr Ser Arg Trp Val Ile Thr Ala Ala His Cys Phe Lys Asp Asn
264             35             40             45
266 ctg aac aaa cca tac ctg ttc tct gtg ctg ctg ggg gcc tgg cag ctg      336
267 Leu Asn Lys Pro Tyr Leu Phe Ser Val Leu Leu Gly Ala Trp Gln Leu
268             50             55             60
270 ggg aac cct ggc tct cgg tcc cag aag gtg ggt gtt gcc tgg gtg gag      384
271 Gly Asn Pro Gly Ser Arg Ser Gln Lys Val Gly Val Ala Trp Val Glu
272             65             70             75
274 ccc cac cct gtg tat tcc tgg aag gaa ggt gcc tgt gca gac att gcc      432
275 Pro His Pro Val Tyr Ser Trp Lys Glu Gly Ala Cys Ala Asp Ile Ala
276     80             85             90
278 ctg gtg cgt ctc gag cgc tcc ata cag ttc tca gag cgg gtc ctg ccc      480
279 Leu Val Arg Leu Glu Arg Ser Ile Gln Phe Ser Glu Arg Val Leu Pro
280 95             100             105             110
282 atc tgc cta cct gat gcc tct atc cac ctc cct cca aac acc cac tgc      528
283 Ile Cys Leu Pro Asp Ala Ser Ile His Leu Pro Pro Asn Thr His Cys
284             115             120             125
286 tgg atc tca ggc tgg ggg agc atc caa gat gga gtt ccc ttg ccc cac      576
287 Trp Ile Ser Gly Trp Gly Ser Ile Gln Asp Gly Val Pro Leu Pro His
288             130             135             140
290 cct cag acc ctg cag aag ctg aag gtt cct atc atc gac tcg gaa gtc      624
291 Pro Gln Thr Leu Gln Lys Leu Lys Val Pro Ile Ile Asp Ser Glu Val
292             145             150             155
294 tgc agc cat ctg tac tgg cgg gga gca gga cag gga ccc atc act gag      672
295 Cys Ser His Leu Tyr Trp Arg Gly Ala Gly Gln Gly Pro Ile Thr Glu
296     160             165             170

```

## RAW SEQUENCE LISTING

DATE: 11/02/2001

PATENT APPLICATION: US/09/856,298

TIME: 14:40:19

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11022001\I856298.raw

```

298 gac atg ctg tgt gcc ggc tac ttg gag ggg gag cgg gat gct tgt ctg      720
299 Asp Met Leu Cys Ala Gly Tyr Leu Glu Gly Glu Arg Asp Ala Cys Leu
300 175      180      185      190
302 ggc gac tcc ggg ggc ccc ctc atg tgc cag gtg gac ggc gcc tgg ctg      768
303 Gly Asp Ser Gly Gly Pro Leu Met Cys Gln Val Asp Gly Ala Trp Leu
304      195      200      205
306 ctg gcc ggc atc atc agc tgg ggc gag ggc tgt gcc gag cgc aac agg      816
307 Leu Ala Gly Ile Ile Ser Trp Gly Glu Gly Cys Ala Glu Arg Asn Arg
308      210      215      220
310 ccc ggg gtc tac atc agc ctc tct gcg cac cgc tcc tgg gtg gag aag      864
311 Pro Gly Val Tyr Ile Ser Leu Ser Ala His Arg Ser Trp Val Glu Lys
312      225      230      235
314 atc gtg caa ggg gtg cag ctc cgc ggg cgc gct cag ggg ggt ggg gcc      912
315 Ile Val Gln Gly Val Gln Leu Arg Gly Arg Ala Gln Gly Gly Gly Ala
316      240      245      250
318 ctc agg gca ccg agc cag ggc tct ggg gcc cca gcg ctt ttg tgt ata      960
319 Leu Arg Ala Pro Ser Gln Gly Ser Gly Ala Pro Ala Leu Leu Cys Ile
320 255      260      265      270
322 taaatgttaa tgatttttat aggtatttgt aaccctgccc acatatc      1007
325 <210> SEQ ID NO: 4
326 <211> LENGTH: 319
327 <212> TYPE: PRT
328 <213> ORGANISM: Homo sapiens
330 <400> SEQUENCE: 4
332 Met Val Val Ser Gly Ala Pro Pro Ala Leu Gly Gly Gly Cys Leu Gly
333      -45      -40      -35
336 Thr Phe Thr Ser Leu Leu Leu Leu Ala Ser Thr Ala Ile Leu Asn Ala
337      -30      -25      -20
340 Ala Arg Ile Pro Val Pro Pro Ala Cys Gly Lys Pro Gln Gln Leu Asn
341      -15      -10      -5
344 Arg Val Val Gly Gly Glu Asp Ser Thr Asp Ser Glu Trp Pro Trp Ile
345 -1 1      5      10      15
348 Val Ser Ile Gln Lys Asn Gly Thr His His Cys Ala Gly Ser Leu Leu
349      20      25      30
352 Thr Ser Arg Trp Val Ile Thr Ala Ala His Cys Phe Lys Asp Asn Leu
353      35      40      45
356 Asn Lys Pro Tyr Leu Phe Ser Val Leu Leu Gly Ala Trp Gln Leu Gly
357      50      55      60
360 Asn Pro Gly Ser Arg Ser Gln Lys Val Gly Val Ala Trp Val Glu Pro
361      65      70      75
364 His Pro Val Tyr Ser Trp Lys Glu Gly Ala Cys Ala Asp Ile Ala Leu
365 80      85      90      95
368 Val Arg Leu Glu Arg Ser Ile Gln Phe Ser Glu Arg Val Leu Pro Ile
369      100      105      110
372 Cys Leu Pro Asp Ala Ser Ile His Leu Pro Pro Asn Thr His Cys Trp
373      115      120      125
376 Ile Ser Gly Trp Gly Ser Ile Gln Asp Gly Val Pro Leu Pro His Pro
377      130      135      140
380 Gln Thr Leu Gln Lys Leu Lys Val Pro Ile Ile Asp Ser Glu Val Cys

```

Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/856,298

DATE: 11/02/2001

TIME: 14:40:20

Input Set : A:\sequence listing.txt

Output Set: N:\CRF3\11022001\I856298.raw

L:1779 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:30

L:1804 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31